Numerische Mathematik Bd. 1, S. 38–40 (1959)

Orthogonal polynomials in several variables

By

MORRIS WEISFELD*

The method of A. S. HOUSEHOLDER [2] and E. STIEFEL [3] for orthogonalizing real polynomials in one variable can be generalized to a method for orthogonalizing real polynomials in several variables.

Let D be a set bearing a non-negative measure μ. Given two mappings f and g of D into the reals, their scalar product (f, g) is defined to be $\int_D fg\, d\mu$ where $(fg)(x) = f(x)\, g(x)$ for all $x \in D$. A set of real-valued mappings of D is *orthogonal* if and only if $(f, g) = 0$ for each f and g, $f \neq g$ in the set, and *independent* if and only if no non-trivial finite linear combination of elements in the set is zero almost everywhere. Let $\Phi = \{\varphi_j | j \in J\}$ be an ordered independent square-integrable set of real-valued mappings of D. An *orthogonalization* of Φ is an ordered orthogonal set $\Psi = \{\psi_j | j \in J\}$ of real-valued mappings of D such that for each $i \in J$, φ_i can be written as a finite linear combination of elements of the set $\{\psi_k | k \in J,\, k \leqq i\}$.

We consider the case of D being a subset of R^n, the Cartesian product of n real lines, J being the set of n-tuples of non-negative integers and Φ the set of monomials in the coordinate variables; that is, if $j = (j_1, \ldots, j_n)$ and $x_1, \ldots, x_n$ represent coordinates, $\varphi_j = x_1^{j_1} x_2^{j_2} \ldots x_n^{j_n}$. Define $\sigma(j) = j_1 + \cdots + j_n$. Order J as follows: $i < j$ if and only if $\sigma(i) < \sigma(j)$ or $\sigma(i) = \sigma(j)$ and, for some $k \leqq n$, $i_k + \cdots + i_n < j_k + \cdots + j_n$. This induces an order in Φ.

Define:

$$J_d = \{j | j \in J \quad \text{and} \quad \sigma(j) \leqq d\},$$
$$J_{d,k} = \{j | j \in J, \sigma(j) = d \quad \text{and} \quad j_{k+1} = \cdots = j_n = 0\}\; k = 0, \ldots, n-1; \quad d = 0, 1, \ldots;$$
$$J_{d,n} = \{j | j \in J, \sigma(j) = d\}.$$

The ψ_j are defined as follows:

$$\psi_{(0,\ldots,0)} = \psi_0 = 1. \tag{1}$$

For $j > 0$, the d and k such that $j \in J_{d,k}$ and $j \notin J_{d,k-1}$ are uniquely determined. Define for $j > 0$

$$\hat{j} = (j_1, \ldots, j_k - 1, j_{k+1}, \ldots, j_n). \tag{2}$$

Then define

$$\psi_j = x_k \psi_{\hat{j}} - \sum \alpha_j^m \psi_m \tag{3}$$

where the sum is over all $m < j$ such that $\sigma(j) - \sigma(m) \leqq 2$ and

$$\alpha_j^m = \frac{(x_k \psi_{\hat{j}}, \psi_m)}{(\psi_m, \psi_m)}. \tag{4}$$

* Now at Shell Development Company, Emeryville 8, California.

Since each ψ_j can be written in the form

$$\varphi_j + \sum_{n<i} \beta_j^n \varphi_n$$

the independence of Φ insures that $(\psi_j, \psi_j) \neq 0$ and that the mapping $j \to \psi_j$ induces an order in ψ. The choice of α_j^m insures the orthogonality of the set $\{\psi_n | n \in J, n \leqq j \text{ and } \sigma(j) - \sigma(n) \leqq 2\}$.

We claim: *Let $m \in J_{d,k}$. If $i \geqq k$ then $x_i \psi_m$ is a linear combination of elements of $\{\psi_j | j \in J_d \cup J_{d+1,i}\}$. If $i < k$ then $x_i \psi_m$ is a linear combination of elements of $\{\psi_j | j \in J_d \cup J_{d+1,k}\}$.*

To establish the claim suppose $i \geqq k$. Let $n = (m_1, \ldots, m_{i-1}, m_i + 1, m_{i+1}, \ldots, m_n)$. Then $\hat{n} = m$ and by (3) $x_i \psi_m = \psi_n + \sum \alpha_n^p \psi_p$. Since $p < n$ and $n \notin J_{d+1,i}$, the first assertion is proved.

Now suppose $i < k$. The assertion is vacuously true for $j = (0, \ldots, 0)$. Suppose we have proved it for all $n < m$. There is an $r \leqq k$ such that $m \in J_{d,r}$ and $m \notin J_{d,r-1}$. If $r \leqq i$, we use the first assertion. Hence assume $i < r$. By (3) we have

$$x_i \psi_m = x_i x_r \psi_{\hat{m}} - \sum \alpha_r^n x_i \psi_n . \tag{5}$$

The terms $x_i \psi_n$ are, by the induction hypothesis, a linear combination of elements of $\{\psi_j | j \in J_d \cup J_{d+1,r}\}$. $x_i \psi_{\hat{m}}$ is, by the induction hypothesis a linear combination of $\{\psi_j | j \in J_{d-1} \cup J_{d,r}\}$ and x_r times elements of the latter set are by the first assertion linear combinations of elements of $\{\psi_j | j \in J_d \cup J_{d+1,r}\}$. This proves the second assertion.

Suppose we have shown $\{\psi_n | n \in J, n < j\}$ is orthogonal. To show ψ_j is orthogonal to all the elements of this set it remains to show $(\psi_j, \psi_n) = 0$ for $n < j$ such that $\sigma(j) - \sigma(n) > 2$. We have by (3)

$$(\psi_j, \psi_n) = (x_k \psi_{\hat{j}}, \psi_n) - \sum \alpha_j^m (\psi_m, \psi_n) .$$

Since $\sigma(m) \neq \sigma(n)$ and $m < j$, $(\psi_m, \psi_n) = 0$ for all terms in the sum, so that

$$(\psi_j, \psi_n) = (x_k \psi_{\hat{j}}, \psi_n) = (\psi_{\hat{j}}, x_k \psi_n) .$$

Since $\sigma(\hat{j}) = d - 1$ and $x_k \psi_n$ is a linear combination of elements of $\{\psi_j | j \in J_{d-2}\}$, this scalar product is also zero. Hence $\{\psi_n | n \in J, n \leqq j\}$ is orthogonal.

It is clear that each φ_i can be written as a linear combination of the elements of $\{\psi_n | n \leqq i\}$. Hence our construction yields an orthogonalization of Φ.

A useful observation is the following: *Suppose D is a product measure space $D_1 \times D_2$ endowed with measure $\mu_1 \times \mu_2$. Let $\{f_i | i \in I\}$ and $\{g_j | j \in J\}$ be ordered independent sets of mappings of D_1 and D_2 respectively into the reals, and let $\{h_i | i \in I\}$ and $\{k_j | j \in J\}$ be their respective orthogonalizations. Let $\{f_i g_j | (i, j) \in I \times J\}$ be a set of mappings of D into the reals, ordered by $f_{i_1} g_{j_1} < f_{i_2} g_{j_2}$ if and only if $g_{j_1} < g_{j_2}$ or $g_{j_1} = g_{j_2}$ and $f_{i_1} < f_{i_2}$. Then the last set of mappings is independent, and the set $\{h_i k_j | (i, j) \in I \times J\}$ is an orthogonalization of it.*

For suppose $\sum c_{ij} f_i g_j = 0$, $c_{ik} \neq 0$. Choose $y \in D_2$ so that $g_k(y) \neq 0$. Then $\sum c_{ij} g_j(y) f_i = 0$, violating the independence of $\{f_i | i \in I\}$. Furthermore,

$$(h_{i_1} k_{j_1}, h_{i_2} k_{j_2}) = \int_{D_1} h_{i_1} h_{i_2} d\mu_1 \int_{D_2} k_{j_1} k_{j_2} d\mu_2 = 0 \quad \text{if} \quad (i_1, j_1) \neq (i_2, j_2) .$$

Clearly $f_i g_j$ can be written as a linear combination of elements of

$$\{h_r k_s | (r, s) \in I \times J, \quad (r, s) \leqq (i, j)\} .$$

References

[1] FORSYTHE, G. E.: Generation and use of orthogonal polynomials for data-fitting with a digital computer. J. Soc. Ind. and Appl. Math. 5, No. 2, 74—88 (1957).
[2] HOUSEHOLDER, A. S.: Principles of numerical analysis, p. 221. New York-Toronto-London: McGraw-Hill 1953. 274 pp.
[3] STIEFEL, E. L.: Kernel polynomials in linear algebra and their numerical applications, pp. 1—22 of *Further contributions to the solution of simultaneous linear equations and the deter mination of eigen values*, National Bureau of Standards, Applied Math. Ser. 49, Washington, Government Printing Office, 1958.

San Francisco, California

(Received May 29, 1958)